

Identifying Greek History and Biography

How do we identify the most representative samples for the Greek historical and biographical genres? The answer to this question is complicated by debates revolving around several (especially) early biographical writings. I will argue, however, that careful attention to diachronic issues may help better situate this discussion. I also insist upon the importance of corpus design criteria in the process of corpus compilation. Corpus linguistics provides a powerful resource to the contemporary genre critic. But if neglected, it leaves the interpreter vulnerable to compromised data and, by extension, unstable conclusions. After briefly outlining some of the salient principles of corpus linguistics and their application to New Testament genre study, this chapter will identify a corpus of Greek historical and biographical texts that will provide a basis for typological comparisons with Luke-Acts. In this way, the present chapter sets up the remainder of the book, which will compare the Greek history and the βίος to Luke-Acts *via* the eight genre agnation scales developed in the prior chapter.

1 What is a ‘Representative’ Corpus and Why is it Important?

A great deal of attention is given in contemporary genre study to corpus compilation. Graeme Kennedy urges that the ‘purpose of sampling adequately’ is ‘so that generalizations can be made readily and validly about the whole.’¹ Or as Matthew Brook O’Donnell, in his work on *Corpus Linguistics and the Greek of the New Testament*, insists: when compiling Hellenistic texts in order to make comparative assessments relative to the New Testament on issues like genre,

The Hellenistic Greek linguist must work primarily with the extant body of documents from the Hellenistic period ... The limited and accidental nature of the extant Hellenistic literature ... requires the careful attention to sampling and compilation criteria that built earlier corpora ... In order to make generalizations and test theories concerning the nature of

¹ G.D. Kennedy, *An Introduction to Corpus Linguistics* (New York: Routledge, 2016), 74.

language, these texts must be classified and grouped in a manner that allows for representative statements to be made.²

In other words, a carefully constructed (i.e. representative) corpus is more likely to yield reliable results, and vice versa.

In one recent treatment, exploring the relationship of corpus linguistics to genre theory, Chengyu Fang and Jing Cao note, 'By and large in the past 50 years, the representative function of the corpus has been thoroughly understood and appreciated' in genre studies.³ When creating a representative corpus, two theoretical issues confront the interpreter of ancient genres: corpus design and compilation.⁴ As Kennedy puts it, 'Issues in corpus design and compilation are fundamentally concerned with validity and reliability of research, based on a particular corpus, including whether that corpus can serve the purpose for which it was intended'.⁵ And this raises a range of further issues: (1) static vs. dynamic corpora; (2) the degree to which the corpus is representative of a language or genre; and (3) corpus and sample text size.⁶ These criteria assess or provide diagnostics for corpus design in terms of both corpus *quality* and *quantity*.

Corpus Quality

1. *Static vs. Dynamic Corpora*. Static corpora seek to establish a very tight group of related texts as a kind of snapshot of the language. These are synchronic in nature. Dynamic or monitor corpora, are so called because they monitor changes in a genre over time. These corpora, then, are diachronic in nature.
2. *Representativeness and Balance*. Kennedy raises the question: representative of what? This depends upon the research focus. If one seeks to investigate grammatical features in Greek, they will need a representative corpus of the Greek language, including its many genres.⁷ When analyzing a more specific structure such as genre, one must ensure that the corpus is as 'thoroughly representative' of 'a particular genre, subject field or topic' as possible.⁸ 'The particular year or years from which

2 M.B. O'Donnell, *Corpus Linguistics and the Greek of the New Testament* (NTM 6; Sheffield: Sheffield Phoenix Press, 2005), 137.

3 C. Fang and J. Cao, *Text Genres and Registers: The Computation of Linguistic Features* (Berlin: Springer, 2015), 6.

4 Kennedy, *Introduction*, 60.

5 Kennedy, *Introduction*, 60.

6 Kennedy, *Introduction*, 60–69.

7 See O'Donnell, *Corpus Linguistics*, 135–38.

8 Kennedy, *Introduction*, 62.